

Molecular cloning and functional analysis of Chinese bayberry MrSPL4 that enhances growth and flowering in transgenic tobacco

Xiangqi Wu

Zhejiang Normal University

Shuwen Zhang

Zhejiang Academy of Agricultural Sciences

Zheping Yu

Zhejiang Academy of Agricultural Sciences

Li Sun

Zhejiang Academy of Agricultural Sciences

Senmiao Liang

Zhejiang Academy of Agricultural Sciences

Xiliang Zheng

Zhejiang Academy of Agricultural Sciences

Xingjiang Qi (✉ qxj@zaas.ac.cn)

Zhejiang Normal University

Haiying Ren

Zhejiang Academy of Agricultural Sciences

Article

Keywords: Chinese bayberry, Growth and flowering, MrSPL4, Gene function, qRT-PCR

Posted Date: October 12th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2055794/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Chinese bayberry (*Myrica rubra*) is an important tree in South China, with its fruit being of nutritional and high economic value. The early maturing varieties of Chinese bayberry are characterized by early flowering. In this study, early ripening (ZJ), medium ripening (BQ) and late ripening (DK) varieties were used as test materials, while young leaves in the floral bud morphological differentiation period were selected for transcriptome sequencing. A total of 4,538 differentially expressed genes were detected. Based on clustering analysis and comparisons with genes reportedly related to flowering in *Arabidopsis thaliana*, 25 homologous genes were identified. Of these, one was down-regulated in DK but up-regulated in ZJ and BQ, with the results also reflected in the relative expression determined by qRT-PCR. The gene contained the SBP domain and the target site of miR156, named *MrSPL4*. Primers were designed to clone the gene whose total length and the length of the CDS sequence were 1,664 bp and 555 bp respectively. An overexpression vector (35S::35S::*MrSPL4*-pCambia2301-KY) was then constructed and successfully transfected into tobacco to obtain *MrSPL4*-positive plants. In this case, the relative expression of *MrSPL4* was up regulated by 3,862.0-5,938.4 times. The height of positive plants was also significantly higher than that of wild-type (WT) plants, with the bud stage occurring 12 days earlier. Altogether, the results indicated that the gene influenced growth and flowering, which provided a theoretical basis for the regulation of flowering in Chinese bayberry.

Introduction

Myrica rubra (Lour.) Sieb. et Zucc., of the family *Myricaceae*, is a native, economically important tree in South China where it is particularly concentrated in the south of the Yangtze River Basin¹. Its fruit, Chinese bayberry, is not only soft, juicy and rich in flavor, but in addition to ecological benefits, it also has medical uses. Chinese bayberry is widely favored by consumers, especially in the Zhejiang Province where its fame has, to a large extent, promoted the healthy development of the Chinese bayberry industry to drive the economic development of planting areas^{2,3}.

The maturation period of the existing main varieties of *Myrica rubra* is about 15 days and occurs in the middle to late June. In addition, the subsequent ripening period is very short and coincides with the plum rain season in the South, which causes great harm and serious economic losses. In this context, early maturing varieties can effectively avoid the influence of plum rain, lengthen the maturity period, reduce the market pressure caused by concentrated maturation and significantly improve the economic benefits of cultivation. Therefore, the development of early maturing germplasm as well as the cultivation of new varieties with characteristics of early maturation have become important for the sustainable development of this industry. After years of observation and studies on the development of different flower buds, it was found that early flowering is an important phenotype related to early ripening. Hence, identifying flowering genes and elucidating their mechanism of action can be useful to regulate the ripening stage and create new germplasm with the improved characteristics. In this context, the SQUAMOSA promoter-binding protein-like (*SPL*) gene family is known to be important for the regulation of plant flowering, but its role in the flowering process of Chinese bayberry is yet to be reported.

The *SPL* gene family, also known as SBP protein, is a unique type of transcription factor in green plants whose members have a highly conserved SBP domain⁴. The latter, which is about 80 amino acid residues in length, contains two zinc finger structures (Cys-Cys-His-Cys; Cys-Cys-Cys-His or Cys-Cys-Cys-Cys) as well as a nuclear localization signal (NLS) located at the C-terminal. Most *SPL* genes also contain highly conserved microRNA156/157 (miR156/miR157) targeting sites that regulate more complex physiological processes^{5,6}. *SPLs* were first identified from *Antirrhinum majus* but with the rise of plant genomics, they have been isolated, identified and analyzed from a number of other plants, including *Oryza sativa*⁷, *Arabidopsis thaliana*⁸, *Solanum lycopersicum*⁹, *Malus × domestica* Borkh.¹⁰, *Vitis Vinifera*¹¹, peony (*Paeonia Suffruticosa*)¹², *Fragaria Vesca*¹³, *Citrus sinensis*¹⁴. Currently, *SPLs* are considered to be key genes that regulate biological processes in plants, especially since they show variations in their functions. For instance, these genes can regulate the flowering process of plants^{15,16}, coordinate plant root, stem and leaf development^{17–20}, influence response to biotic and abiotic stress^{21,22} as well as participate in secondary metabolic processes⁴.

In recent years, other functions of *SPL* genes in plants have also been widely studied. For example, in *Arabidopsis thaliana*, *SPL10* was found to be highly expressed in plant leaf and root tissues, resulting in earlier flowering, narrower leaf shapes, smaller and fewer rosette leaves as well as reduced root length and root number by binding to the *AGL79* promoter²³. In addition, the expression of *AtSPL9* and *AtSPL10* in leaf primordia was also reported to affect the differentiation of apical meristem into leaves²⁴, while in leaf tissues, *SPL2* could control floral organs, long silique development and plant fertility by activating *AS2*²⁵. In rice, the GO function analysis of differentially expressed genes in blade leaves of *SPL4* mutant rice showed that *OsSPL4* gene mutations affected protein phosphorylation as well as the binding of iron ions in rice leaves, maintaining the normal plant type of rice²⁶. In pea (*Pisum sativum* L.), *PsSPL3a/3c* was found to be mainly expressed at the transcriptional level in leaves, hence indicating its possible involvement in leaf phase transition in the pea aging pathway²⁷. In maize (*Zea Mays*), *SPL4* plays an important role in bract development and meristem establishment²⁸, while *SPL10/14/26* not only regulates the expression of *ZmWOX3A* and auxin related genes but is also involved in the development of epidermal hair on maize leaves²⁹. Finally, an analysis of the expression of *MdSBP* genes in apple leaves after different hormone treatments showed that many of the genes responded to different plant hormones, thereby suggesting that *MdSBP* genes could be involved in response to hormone signals during stress or apple development¹⁰.

Therefore, based on the previous genome sequencing of *Myrica rubra*², the latter's *SPL* gene family was identified and analyzed based on bioinformatics methods. This was followed by the cloning of an *SPL* gene and its subsequent heterologous expression in *Nicotiana benthamiana* L. by constructing an overexpression vector to validate the functions of the gene. Altogether, this study is expected to provide a theoretical basis for revealing the regulatory pathway of flowering in Chinese bayberry.

Results

Evaluation of transcriptome data of young leaves from different flowering materials

The transcriptome sequencing of nine samples of young leaves (three biological replicates for each variety) in the floral bud morphological differentiation period was completed, and a total of 59.78 Gb of clean data, with an average GC content of 47.28% and a Q30 base ratio of 93.55%, were obtained. After comparison with the reference genome², the percentage of clean reads aligned to the reference genome was found to be 95.39% (Table 1, the transcriptome data of BQ and DK were uploaded to <https://bigd.big.ac.cn/gsa/browse/CRA008253>, and the datasets of ZJ generated and analysed during the current study are available in the NCBI repository <https://www.ncbi.nlm.nih.gov/sra/PRJNA733585>). The number of differentially expressed genes between the three samples was then compared. In this case, 623 genes were differentially expressed between ZJ and DK, and of these, 476 were up-regulated and 147 were down-regulated. Similarly, 2,343 genes were differentially expressed between ZJ and BQ, with 1,385 and 958 genes being up-regulated and down-regulated respectively. Finally, the number of differentially expressed genes between DK and BQ was 1,572, with 734 and 838 being up-regulated and down-regulated respectively (Table 2).

Table 1 Statistics of sequencing data and alignment results

Samples	Clean reads	Clean bases (bp)	GC Content (%)	≥Q30 (%)	Mapped Reads Ratio (%)
BQ-1	21,403,050	6,390,283,060	47.35	92.68	95.32
BQ-2	20,776,437	6,193,566,414	47.38	93.55	95.82
BQ-3	27,227,360	8,115,985,288	47.30	93.79	95.68
ZJ-1	20,138,974	6,004,597,524	47.07	93.77	94.96
ZJ-2	20,316,870	6,058,667,134	47.77	93.22	95.98
ZJ-3	21,634,935	6,455,567,006	47.39	93.85	96.31
DK-1	23,963,165	7,146,744,068	47.72	93.51	95.52
DK-2	23,797,007	7,089,937,938	46.77	93.99	93.73
DK-3	21,214,319	6,327,563,846	46.78	93.59	95.17
Total/Average	200,472,117	59,782,912,278	47.28	93.55	95.39

Table 2 Differentially expressed genes between pairs of samples

DEG Set	DEG Number	Up-regulated	Down-regulated
BQ Vs DK	1,572	734	838
ZJ Vs BQ	2,343	1,385	958
DK Vs ZJ	623	476	147

KEGG analysis of differentially expressed genes

Through KEGG enrichment analysis, it was found that the differentially expressed genes mainly involved functions such as cellular processes, environmental information processing, genetic information processing, metabolism and organic systems (Fig. 1). The pathways that were significantly enriched in all the three groups included phytohormone signal transduction, sulfur and carbon metabolism, fatty acid, phenylpropionic acid, pyruvic acid, α -linolenic acid metabolism, glycine, serine, threonine, arginine, proline, cyano amino acids, cysteine and methionine metabolism, terpenoid skeleton biology, carotenoid biosynthesis, glutathione and glycerophosphatide metabolism, glycolysis/gluconeogenesis, starch and sucrose, amino sugar and nucleotide sugar metabolism, protein processing in the endoplasmic reticulum. The results indicated that these pathways could participate in the regulation network of Chinese bayberry flowering or other important pathways.

Mining of flowering-related differentially expressed genes

The differentially expressed genes mentioned above were compared with 306 flowering genes reported in *Arabidopsis*³⁰, and 25 genes were found to be homologous in Chinese bayberry (Fig. 2a). In particular, one of the differentially expressed genes, MRNA_003335_1, was down-regulated in DK but up-regulated in ZJ and BQ. This gene also contained the *SBP* domain and belonged to the *SPL* gene family, named *MrSPL4*. The relative expression of *MrSPL4* was therefore verified by qRT-PCR (Fig. 2b), and the results showed that ZJ had the highest expression, followed by BQ, with DK showing the lowest expression level. These results were, in fact, consistent with the expression determined by the transcriptome.

Gene analysis of *SPLs* family in Chinese bayberry

Through the screening of all genes in the reported genome provided in Ren et al. (2019), 17 *SPL* family genes with *SBP* domains were found in the Chinese bayberry genome (Table 3). Of these, 12 genes including *MrSPL4*, contained the target site of miR156 in the CD region. The software MEGA7.0 was then used to analyze the evolution of 17 *SPL* genes in *Myrica rubra* (*MrSPL*), 16 *SPL* genes in *Arabidopsis thaliana* and 46 *SPL* genes in apple (Fig. 3). In this case, it was observed that the *SPL* proteins could be divided into four different groups (I, II, III and IV), with each containing at least one *MrSPL* gene. More specifically, groups I and II contained one *MrSPL* each, group IV contained twelve *MrSPLs* and group III contained three Chinese bayberry *SPL* genes, including *MrSPL4*. *MrSPL4* had the highest homology with the AT1G20980 (*AtSPL14*) gene in *Arabidopsis thaliana*, with previous studies showing that this gene (*AtSPL14*) not only promoted the normal growth and development of *Arabidopsis thaliana*, but also

played a crucial role in the development of flowering as well as the transformation from a vegetative to reproductive growth. Thus, it was speculated that *MrSPL4* could be playing an important role in the flowering process of Chinese bayberry.

Table 3 Gene status of the SPL family in *Myrica rubra* genome

Gene ID	Conserved domain	Target site of miR156
MRNA_003237_1	<i>SBP</i> domain	No
MRNA_013273_1	<i>SBP</i> domain	No
MRNA_013732_2	<i>SBP</i> domain	Yes
<i>MrSPL4</i>	<i>SBP</i> domain	Yes
MRNA_019617_1	<i>SBP</i> domain	No
MRNA_009540_1	<i>SBP</i> domain	Yes
MRNA_018907_1	<i>SBP</i> domain	Yes
MRNA_022003_2	<i>SBP</i> domain	Yes
MRNA_023074_1	<i>SBP</i> domain	Yes
MRNA_023065_2	<i>SBP</i> domain	No
MRNA_000299_1	<i>SBP</i> domain	No
MRNA_007729_2	<i>SBP</i> domain	Yes
MRNA_013669_4	<i>SBP</i> domain	Yes
MRNA_016389_2	<i>SBP</i> domain	Yes
MRNA_001354_1	<i>SBP</i> domain	Yes
MRNA_008983_1	<i>SBP</i> domain	Yes
MRNA_000723_2	<i>SBP</i> domain	Yes

Cloning, structural analysis and construction of overexpression vector of *MrSPL4* gene

MrSPL4 is located at scaffold 10: 3,747,409-3,749,072 bp of the reference genome. Primers were therefore designed based on the reference genome sequence (Table 4), with ORF sequences of the *MrSPL4* gene in ZJ, BQ and DK subsequently obtained by PCR amplification. The electrophoresis results (Fig. 4a) showed that the bands were consistent with the expected amplification product size, hence indicating that the sequence of the coding region of *MrSPL4* was successfully obtained. No differences in the gene sequence were noted between the three samples (Fig. 4b). Therefore, it was inferred that the expression of this gene was different between the different test materials, probably due to the promoter element, but this remained to be experimentally verified. The full length and length of the CD sequence of

MrSPL4 were 1,664 bp and 555 bp respectively. The gene also contained two exons and one intron (Fig. 4c), along with a binding site of *miRNA156* in the CDS1 region (Fig. 4d).

Table 4 Primer sequences

Primer name	Primer sequence (5' to 3')
<i>MrSPL4</i> -F	ATGTCTAGACTCGAGGGATCCATGAATACGGGAAAGGTTGATG
<i>MrSPL4</i> -R	AGCCTGCAGCCATGGGGATCCTTATCTGATCTGGAAATGCTTG
Q- <i>MrSPL4</i> -F	GCATAATGAACGGCGAAGG
Q- <i>MrSPL4</i> -R	ATCTGGAAATGCTTGTAAGTGG
Q- <i>NtActin</i> -F	CTATTCTCCGCTTTGGACTTGGCA
Q- <i>NtActin</i> -R	ACCTGCTGGAAGGTGCTGAGGGAA

The amplified product was first recovered, and the vector was digested with BamHI. The resulting enzymatically digested product was then recombined with the amplified product to construct a plant overexpression vector. The latter was transformed into *E. coli* competent cells DH-5α before identifying the transformed bacterial solution by PCR (Fig. 4e). Results, in this case, showed that bands 1, 3, 5, 6, 7 and 8 were consistent with the expected target fragments. In fact, preliminary results further showed that the *MrSPL4* gene was successfully inserted into the vector to yield six positive transformant colonies. Two of these positive colonies were randomly selected for sequencing, with the results being still consistent. This experiment therefore showed that the overexpression vector of *MrSPL4* gene was successfully constructed and named as 35S::*MrSPL4*-pCambia2301-KY.

Functional validation of *MrSPL4* in tobacco

Verification of *MrSPL4* positive tobacco plants

Agrobacterium-mediated transformation of *Nicotiana benthamiana* with the recombinant plasmid 35S::*MrSPL4*-pCambia2301-KY was performed. The tobacco leaves infected by *Agrobacterium tumefaciens* were directly transferred to a selective medium containing kanamycin (Kan) to induce differentiation and budding (Fig. 5a). After the buds had grown to 2-3cm (Fig. 5b), they were inserted into a rooting medium to induce the formation of roots (Fig. 5c). Once the root system was vigorous, the seedlings were then tempered and transplanted to soil for culture (Fig. 5d) to obtain completely regenerated tobacco with Kan resistance.

Leaf DNA from the resistant regenerated tobacco plants to be tested was used as a template for PCR-based validation. For this purpose, *MrSPL4*-F and *MrSPL4*-R were used as primers, water was set as a blank control, the 35S::*MrSPL4*-pCambia2301-KY expression vector plasmid was used as a positive control and leaf DNA of WT plants was used as a negative control. The results (Fig. 5e) showed that for all resistant regenerated tobacco, the target band was amplified, with its size being consistent with that of

the positive control. In addition, no specific bands were observed in WT tobacco and the blank control. Hence, the results showed that the *MrSPL4* gene was successfully transferred into tobacco.

Relative expression and phenotype of *MrSPL4* in positive tobacco

Three positive tobacco lines (35S::*MrSPL4*) from the T₁ generation were randomly selected to detect the expression level of the *MrSPL4* gene. Results showed that the relative expression level was significantly higher than that of WT tobacco, with the up-regulation multiple being between 3,862.0-5,938.4 (Fig. 6a). This was a clear indication that the gene was overexpressed in transformed tobacco.

The plant heights for 35S::*MrSPL4*-transformed tobacco and WT ones were measured on 34 d, 41 d and 78 d after transplanting and found to be significantly higher for the transformed plants compared with the WT (Fig. 6b). Furthermore, the growth rate was also significantly faster than for the WT. In terms of the budding stages, those of 35S::*MrSPL4*-transformed tobacco and WT plants were 34 days and 46 days after transplanting, respectively. Based on the above, it could be concluded that 35S::*MrSPL4*-transformed tobacco showed characteristics of rapid plant growth and early flowering (Fig. 6c). Therefore, it was speculated that *MrSPL4* gene affected the phenotype of transgenic tobacco to promote plant growth and flowering.

Discussion

As a specific and important transcription factor in plants, the *SPL* gene family has a highly conserved SBP domain which plays an important regulatory role in plant growth and development. Although the *SPL* gene family has been widely isolated and identified in many plants such as *Arabidopsis* and rice, research on its role in *Myrica rubra*, an economically important fruit in South China, has not been reported. Previous studies⁷ have shown that the number of *SPL* gene family members varies in different species, thereby leading to the diversification of gene functions and this was confirmed in the current study. In addition, 17 *MrSPL* gene family members with SBP domains were identified in the genome of *Myrica rubra*, with this number being close to that of *SPL* gene family members in *Arabidopsis*⁸ and tomato³¹, but greatly different from that of *Gossypium hirsutum* L.³² and apple¹⁰. In general, members in the same subgroup are likely to have the same or quite similar functions. For example, *AtSPL2*, *AtSPL10* and *AtSPL11* inhibit root growth, while other members of this group, *CsSPL2* and *CsSPL10*, also participate in the regulation of root development⁴. In the present study, phylogenetic-based analyses showed that *SPLs* could be divided into four groups, with each containing at least one *MrSPL* gene. Since *MrSPL4* was found to be homologous with the *AtSPL14* gene, it was therefore speculated that *MrSPL4* could be playing a similarly important role in plant growth and development, although it is likely that the gene could also have different functions.

Previous studies have found that flowering is an important sign of plant growth and development, and consequently, research on the role of the *SPL* gene family in the regulation of plant flowering has attracted significant interest. For example, *AtSPL3/4/5* participates in the photoperiod and the age

pathway, and as such, it can promote the early flowering of *Arabidopsis* by upregulating the expression of downstream genes³³. Similarly, overexpression of the *EjSPL3/4/5/9* genes in loquat causes transgenic *Arabidopsis thaliana* to exhibit characteristics of early flowering³⁴, while strawberry *FvSPL10-OE* plants were shown to bloom 3–5 days earlier³⁵. Despite the above observations, the functions of *MrSPL4* in the flowering process of Chinese bayberry remains unknown. In order to verify its role, the gene was cloned from Chinese bayberry to yield transgenic tobacco overexpressing *MrSPL4*. In this study, the relative expression of *MrSPL4* positive tobacco plants was significantly increased by 3,862.0–5,938.4 times compared with WT under long sunshine conditions. Moreover, the plant heights of transformed tobacco plants were significantly higher than WT tobacco, with the budding period also occurring 12 days earlier. This indicated that the *MrSPL4* gene responded to the flowering process of transgenic tobacco, showing early flowering and increased plant height. In addition, the current study found that the sequence of the *MrSPL4* gene in different Chinese bayberry varieties had no differences, although its expression level did differ in different Chinese bayberry varieties. It was speculated that these differences could be linked to promoter elements but this would need follow-up experiments for validation.

To sum up, 17 members of the *SPL* gene family with SBP domains were identified in *Myrica rubra*. Of these, the *MrSPL4* gene was isolated, cloned and verified in tobacco. The results showed that *MrSPL4* could regulate the flowering process of plants, accelerate their growth and endow the plants with early flowering phenotypes, thus supporting the view that this gene exerted multiple regulatory functions on plant growth and development. These results also provide a basis for further elucidating *MrSPL4*'s regulatory mechanism for flowering in *Myrica rubra* in order to achieve genetic improvement and gene breeding of this plant in the future. Therefore, the *MrSPL4* gene needs to be further studied, especially with regards to its promoter region.

Materials And Methods

Material information

By referring to the expression of genes related to male and female flowering as described by Jia et al. (2019), three experimental materials with different flowering stages were selected from the Lanxi International Chinese bayberry Research Center in November 2019 (period during which floral buds can be morphologically differentiated). These included the early maturing variety “Zaojia” (ZJ), the medium maturing variety “Biqizhong” (BQ) and the late maturing variety “Dongkui” (DK) (Table 5). Each variety was sampled in triplicates, and for transcriptome sequencing, young leaves (not unfolded) were taken from annual branches facing south and at 1 m above the ground.

Table 5 Phenological periods of different test materials

Test material	No.	Flower bud formation time	First flowering period (month-day)	Maturity period (month-day)
Zaojia	ZJ	Mid-November	April 1st	June 6th
Biqizhong	BQ	Mid-December	April 4th	June 15th
Dongkui	DK	Mid-January	April 10th	June 25th

Transcriptome sequencing and screening of differentially expressed gene

A polysaccharide polyphenol RNA extraction kit was used for extracting total RNA from the samples (TIANGEN, Beijing), and after RNA detection, Biomarker Biotechnology Co., Ltd. was commissioned to carry out the transcriptome sequencing. For this purpose, magnetic beads with Oligo (dT) were used to enrich the total RNA of the samples before fragmenting the mRNA with the fragmentation buffer. The first cDNA strand was then synthesized with random hexamers using mRNA as template, and this was followed by the synthesis of the second cDNA strand by adding buffer, dNTPs, RNase H and DNA polymerase I. After purification with the QiaQuick PCR kit and elution with EB buffer, end repair was performed, poly (A) tails were added and sequencing adaptors were connected. Appropriate fragments were then selected by gel electrophoresis prior to PCR-based amplification. The resulting libraries were eventually sequenced on an Illumina HiSeq4000.

After gene splicing, protein sequences were aligned with those from eight public databases (COG, GO, KEGG, KOG, Pfam, Swissport, eggNOG and Nr) using a threshold of $e \leq e^{-10}$. The BLAST algorithm was then used for sequence similarity comparison, with the resulting sequence similarities subsequently used for functional annotations. Relative gene expression was assessed based on RPKM (Reads Per Kilobase of exon model per Million mapped reads) where larger RPKM values were indicative of higher expression levels³⁶.

Differentially expressed genes were screened by the false discovery rate (FDR)³⁷, with a $|\log_2$ fold change ≥ 2 and an FDR < 0.5 selected as thresholds for a gene to be considered as being differentially expressed.

SPL gene family analysis

The *SPLs* of Chinese bayberry were isolated and identified by tBLASTn analysis of AtSPL amino acid sequences obtained from the genomic data of Chinese bayberry². The Chinese bayberry *SPLs* and target sites of miRNA156 were then predicted and confirmed using Genscan Web (<http://genes.mit.edu/GENSCAN.html>) as well as the BLASTx algorithm (<http://www.ncbi.nlm.nih.gov/BLAST>). After obtaining the nucleotide and amino acid sequences of 16 Arabidopsis and 46 apple *SPL* family genes from the Plant Transcription Factor Database (PlnTFDB3.0) (<http://plntfdb.bio.uni-potsdam.de/v3.0/>), phylogenetic trees were also constructed using the NJ method in MEGA 7.0, along with full-length protein sequences and the test parameter (bootstrap) set to 1000. The exon and intron structures of Chinese

bayberry *SPL* genes were obtained by Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/index.php>).

Strains and vectors

Escherichia coli competent cell DH-5 α (Shanghai Jinchao Technology Development Co., LTD.), *Agrobacterium tumefaciens* strain GV3101 and pCambia2301-KY vectors (Shanghai Kaiyi Biotechnology Co., LTD.) were the main requirements of the study.

Primer design and gene cloning

Using the genome sequence of Chinese bayberry², specific primers for both sides of the open reading framework (ORF) of the target gene were designed with Primer Premier 5.0 software for gene cloning. Total RNA extraction was also performed on healthy Chinese bayberry leaves using the modified cetyl trimethyl ammonium bromide (CTAB) method, with the extracted RNA acting as template to synthesize cDNA according to the instructions of the HiScript 1st Strand cDNA Synthesis Kit (Vazyme). This was followed by PCR amplification with the Phanta Max ultra-fidelity DNA polymerase (Vazyme), using the cDNA as template. In this case, each reaction consisted of the following component: 1 μ L of Phanta Max super-Fidelity DNA Polymease, 2 μ L of cDNA, 2 μ L each of both forward and reverse primers, 25 μ L of 2 $\times$ Phanta Max Buffer, 1 μ L of dNTP Mix and ddH₂O (for making up the volume to 50 μ L), while the PCR procedure involved an annealing temperature of 49 °C and an extension rate of 1 kb/min, carried out for 39 cycles. Other operations shall follow the product instructions of Vazyme Company. The amplified products were finally detected on 1.5% agarose gel, before being sent to the company for sequencing to verify the accuracy of cloning results.

Construction of an overexpression vector and *Agrobacterium* transformation

The restriction enzyme BamHI (Takara Company) was first used to linearize the vector before extracting the pCambia2301-KY plasmid for digestion with the same enzyme. The overexpression vector was then constructed with Vazyme recombinant enzyme at 37 °C by using the following components: 2 μ L of 5 $\times$ CE II Buffer, 1 μ L of Exnase II, 4 μ L of linearized carrier, 1 μ L of insert fragment and ddH₂O (to make up the volume to 10 μ L). After 30 min of reaction, the vector was placed on ice for cooling. The cells were then transfected into competent *E. coli* DH-5 α cells and cultured in LB medium containing 50 mg/L Kan. This enabled the selection and subsequent culture of resistant colonies for the positive detection of the gene by PCR. The amplified products were finally sent for sequencing. The positive transformer colony plasmid was extracted and transfected into *Agrobacterium tumefaciens* GV3101 and sterile glycerol was added to preserve the bacteria at -80 °C until required for the next transfection.

***Agrobacterium tumefaciens*-mediated transfection of tobacco**

Tobacco Benn was selected for this set of transformation experiment. WT tobacco was infected with *Agrobacterium* carrying recombinant vector plasmids of target genes using the leaf disk method. After

four times of continuous screening/subculture, resistant buds were eventually recovered and transferred to a rooting medium to induce roots. Once the root system was vigorous, healthy and completely regenerated plants were transplanted to the soil (nutrient soil-vermiculite ratio was 1:1 or 2:1) where they were maintained until the T₀ generation for seed collection.

Collected seeds were sterilized with 70% ethanol, 30% sodium hypochlorite or 40% of 84 disinfectant and rinsed with sterile water 5-6 times. Seeds were then added to 1/2 MS solid selective medium containing 80 mg/L of Kan, and vernalized at 4°C for 2 days to break dormancy. They were subsequently cultured in a light incubator of the laboratory of Zhejiang Academy of Agricultural Sciences (light 28°C, 16 h, Darkness 25°C, 8 h, humidity 50%-70%). After about a week, the seeds were transferred to the soil to maintain grow. The leaves of the transgenic resistant plants and the wild-type ones of the T₁ generation were randomly sampled and stored at -80°C after being frozen in liquid nitrogen.

Determination of relative gene expression

Transgenic positive plants of T₁ generation were obtained through screening with 80 mg/L Kan, 1/2 MS solid selective medium and PCR. The leaves of grown plants were collected, and total RNA was extracted with the RNA simple Total RNA Kit (TIANGEN) after quick-freezing in liquid nitrogen. In addition, synthesis reactions were also performed in 10-μL reaction volumes with the first Strand cDNA synthesis kit. For this purpose, the following components were used as required by the FastFire qPCR PreMix (SYBR Green) Kit (TIANGEN): 5 μL of 2×FastFire qPCR PreMix, 1 μL of forward primer and reverse primers (10 μm) and 1 μL of cDNA template. The reaction was performed on a Light Cycler 96 real-time PCR instrument under the following conditions: 95°C for 60 s, followed by 45 cycles, each at 95°C for 5 s, 63°C for 10 s and 72°C 15 s. Three technical replicates were set for each sample. Quantitative primers were designed according to gene sequences, with *Ntactin-F/R* selected as the reference gene³⁸, and WT tobacco acting as the control to determine the relative expression of target genes. The $2^{-\Delta\Delta C_t}$ method was used to process the data³⁹, while the IBM SPSS Statistics 22 and Origin 2022 were used for statistical analysis and plotting respectively.

Declarations

Ethics statement

The collection of plant material comply with relevant institutional, national, and international guidelines and legislation.

Data availability

All data generated or analyzed during this study are included in this published article and supplementary material.

Acknowledgements

This work was funded from the Special breeding program for new varieties in Zhejiang (2021C02066-2) and the Key R & D in Zhejiang (2021C02009) projects. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

W. X., Y. Z. and Z. S. performed the experiments. Q. X. assisted with design of the project. S. L., L. S., Z. X. and R. H. assisted with the primary data analysis. W. X. and Z. S. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

References

1. Zhang, S. W., *et al.* An overview of the nutritional value, health properties, and future challenges of Chinese bayberry. *PeerJ*. **10**, e13070. <https://doi.org/10.7717/peerj.13070> (2022).
2. Ren, H. Y., *et al.* Genome sequencing provides insights into the evolution and antioxidant activity of Chinese bayberry. *BMC Genomics*. **20**(1), 458. <https://doi.org/10.1186/s12864-019-5818-7> (2019).
3. Jia, H. M., *et al.* The red bayberry genome and genetic basis of sex determination. *Plant Biotechnol J*. **17**(2), 397-409. <https://doi.org/10.1111/pbi.12985> (2019).
4. Yang, J., Guo, Z. L., Wang, W. T., Cao, X. Y., Yang, X. Z. Genome-wide characterization of SPL gene family in *Codonopsis pilosula* reveals the functions of *CpSPL2* and *CpSPL10* in promoting the accumulation of secondary metabolites and growth of *C. pilosula* hairy root. *Genes*. **12**(10), 1588. <https://doi.org/10.3390/genes12101588> (2021).
5. Birkenbihl, R. P., Jach, G., Saedler, H. Huijser, P. Functional dissection of the plant-specific SBP-domain: overlap of the DNA binding and nuclear localization domains. *J. Mol. Biol.* **352**(3), 585-596. <https://doi.org/10.1016/j.jmb.2005.07.013> (2005).
6. Li, J., Gao, X. Y., Zhang, X., Liu, C. Dynamic expansion and functional evolutionary profiles of plant conservative gene family SBP-box in twenty two flowering plants and the origin of miR156. *Biomolecules*. **10**(5), 757. <https://doi.org/10.3390/blom10050757> (2020).
7. Yang, Z. F., *et al.* Comparative study of SBP-box gene family in *Arabidopsis* and rice. *Gene*. **407**(1-2), 1-11. <https://doi.org/10.1016/j.gene.2007.02.034> (2008).
8. Wu, G., *et al.* The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell*. **138**(4), 750-759. <https://doi.org/10.1016/j.cell.2009.06.031> (2009).
9. Salinas, M., Xing, S. P., Höhmann, S., Berndtgen R. & Huijser, P. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. *Planta*. **235**, 1171-1184. <https://doi.org/10.1007/s00425-011-1565-y> (2012).

10. Li, J., *et al.* Genome-wide identification and analysis of the SBP-box family genes in apple(*Malus × domestica* Borkh.). *Plant Physiol Biochem.* **70**, 100-114. <https://doi.org/10.1016/j.plaphy.2013.05.021> (2013).
11. Hou, H. M., *et al.* Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. *PloS One.* **8**(3), e59358. <https://doi.org/10.1371/journal.pone.0059358> (2013).
12. Zhu, F. Y., *et al.* Morphological and physiological changes, and the functional analysis of *PdSPL9* in the juvenile-to-adult phase transition of *paeonia delavayi*. *Plant Cell Tiss Organ Cult.* **133**(3), 325-337. <https://doi.org/10.1007/s11240-018-1384-y> (2018).
13. Xiong, J. S., *et al.* Genome-wide identification and expression analysis of the *SPL* gene family in woodland strawberry *Fragaria vesca*. *Genome.* **61**(9), 1-9. <https://doi.org/10.1139/gen-2018-0014> (2018).
14. Liu, M. Y., Wu, X. M., Long, J. M., Guo, W. W. Genomic characterization of miR156 and *SQUAMOSA promoter binding protein-like* genes in sweet orange (*Citrus sinensis*). *Plant Cell Tiss Organ Cult.* **130**(1), 103-116. <https://doi.org/10.1007/s11240-017-1207-6> (2017).
15. Lei, M., *et al.* Constitutive expression of *Aechmea fasciata SPL 14* (*AfSPL 14*) accelerates flowering and changes the plant architecture in *Arabidopsis*. *Int J Mol Sci.* **19**(7), 2085. <https://doi.org/10.3390/ijms19072085> (2018).
16. Guo, J. Q., *et al.* *SPL7* and *SPL8* represent a novel flowering regulation mechanism in switchgrass. *New Phytol.* **222**(3), 1610-1623. <https://doi.org/10.1111/nph.15712> (2019).
17. Yu, N., Niu, Q. W., Ng K. H., Chua, N. H. The role of miR156/SPLs modules in *Arabidopsis* lateral root development. *Plant J.* **83**(4), 673-685. <https://doi.org/10.1111/tpj.12919> (2015).
18. Li, B. B., *et al.* Genome-wide identification, gene cloning, subcellular location and expression analysis of *SPL* gene family in *P. granatum L.* *BMC Plant Biol.* **21**(1), 400. <https://doi.org/10.1186/s12870-03171-7> (2021).
19. Wang, L., *et al.* A spatiotemporally regulated transcriptional complex underlies heteroblastic development of leaf hairs in *Arabidopsis thaliana*. *EMBO J.* **38**(8), e100063. <https://doi.org/10.15252/embj.2018100063> (2019).
20. Wang, J. W., *et al.* MiRNA control of vegetative phase change in trees. *PLoS Genet.* **13**(7), e0200762. <https://doi.org/10.1371/journal.pgen.1002012> (2018).
21. Feyissa, B.A., Arshad, M., Gruber, M. Y., Kohalmi, S. E., Hannoufa A. The interplay between *miR156/SPL 13* and *DFR/WD40-1* regulate drought tolerance in alfalfa. *BMC Plant Biol.* **19**(1), 434. <https://doi.org/10.1186/s12870-019-2059-5> (2019).
22. Ning, K., *et al.* Molecular characterization and expression analysis of the *SPL* gene family with *BpSPL9* transgenic lines found to confer tolerance to abiotic stress in *Betula platyphylla Suk.* *Plant Cell Tiss Organ Cult.* **130**(3), 469-481. <https://doi.org/10.1007/s11240-017-1226-3> (2017).
23. Gao, R. M., Wang, Y., Gruber, M. Y., Hannoufa, A. MiR156/SPL10 modulates lateral root development, branching and leaf morphology in *Arabidopsis* by silencing *AGAMOUS-LIKE 79*. *Front Plant Sci.* **8**,

2226. <https://doi.org/10.3389/fpls.2017.02226> (2018).
24. Wang, J. W., Schwab, R., Czech, B., Mica, E., Weigel, D. Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*. *Plant Cell*. **20**(5), 1231-1243. <https://doi.org/10.1105/tpc.108.058180> (2008).
25. Wang, Z. S., Wang, Y., Kohalmi, S. E., Amyot, L., Hannoufa, A. SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2 controls floral organ development and plant fertility by activating *ASYMMETRIC LEAVES 2* in *Arabidopsis thaliana*. *Plant Mol Biol*. **92**, 661–674. <https://doi.org/10.1007/s11103-016-0536-x> (2016).
26. Hu, J. H., *et al.* The elite alleles of *OsSPL4* regulate grain size and increase grain yield in rice. *Rice*. **14**(1), 90. <https://doi.org/10.1186/s12284-021-00531-7> (2021).
27. Vander Schoor, J. K., Hecht, V., Aubert, G., Burstin, J., Weller, J. L. Defining the components of the miRNA156-SPL-miR172 aging pathway in pea and their expression relative to changes in leaf morphology. *Plant Gene*. **30**, 100354. <https://doi.org/10.1016/j.plgene.2022.100354> (2022).
28. Chuck, G., Whipple, C., Jackson, D., Hake, S. The maize SBP-box transcription factor encoded by tasselsheath4 regulates bract development and the establishment of meristem boundaries. *Development*. **137**(8), 1243-1250. <https://doi.org/10.1242/dev.048348> (2010).
29. Kong, D.X., *et al.* ZmSPL10/14/26 are required for epidermal hair cell fate specification on maize leaf. *New Phytol*. **230**(4), 1533-1549. <https://doi.org/10.1111/nph.17293> (2021).
30. Bouché, F., Lobet, G., Tocquin, P., Périlleux, C. FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res*. **44**(D1), D1167-1171. <https://doi.org/10.1093/nar/gkv1054> (2016).
31. Cui, L., *et al.* MiR156a-targeted SBP-Box transcription factor SISPL13 regulates inflorescence morphogenesis by directly activating *SFT* in tomato. *Plant Biotechnology J*. **18**, 1670-1682. <https://doi.org/10.1111/pbi.13331> (2020).
32. Cai, C. P., Guo, W. Z., Zhang, B. H. Genome-wide identification and characterization of SPL transcription factor family and their evolution and expression profiling analysis in cotton. *Sci Rep*. **8**(1), 762. <https://doi.org/10.1038/s41598-017-18673-4> (2018).
33. Hyun, Y., *et al.* Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the *Arabidopsis* shoot meristem. *Dev Cell*. **37**(3), 254-266. <https://doi.org/10.1016/j.devcel.2016.04001> (2016).
34. Jiang, Y. Y., *et al.* The role of *EjSPL3*, *EjSPL4*, *EjSPL5*, and *EjSPL9* in regulating flowering in loquat (*Eriobotrya japonica* Lindl.). *Int J Mol Sci*. **21**(1), 248. <https://doi.org/10.3390/ijms21010248> (2019).
35. Xiong, J. S., *et al.* Molecular cloning and characterization of *SQUAMOSA*-promoter binding protein-like gene *FvSPL10* from woodland strawberry (*Fragaria vesca*). *Plants*. **8**(9), 342. <https://doi.org/10.3390/plants8090342> (2019).
36. Trapnell, C., *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. **28**(5), 511-515. <https://doi.org/10.1038/nbt.1621> (2010).

37. Anders, S. Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106. <https://doi.org/10.1038/npre.2010.4282.2> (2010).
38. Zhao, Q., *et al.* *MdbHLH130*, an apple bHLH transcription factor, confers water stress resistance by regulating stomatal closure and ROS homeostasis in transgenic tobacco. *Front Plant Sci.* **11**, 543696. <https://doi.org/10.3389/fpls.2020.543696> (2020).
39. Livak, K. J. Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta(\Delta CT)}$ Method. *Methods.* **25**(4), 402-408. <https://doi.org/10.1006/meth.2001.1262> (2001).

Figures

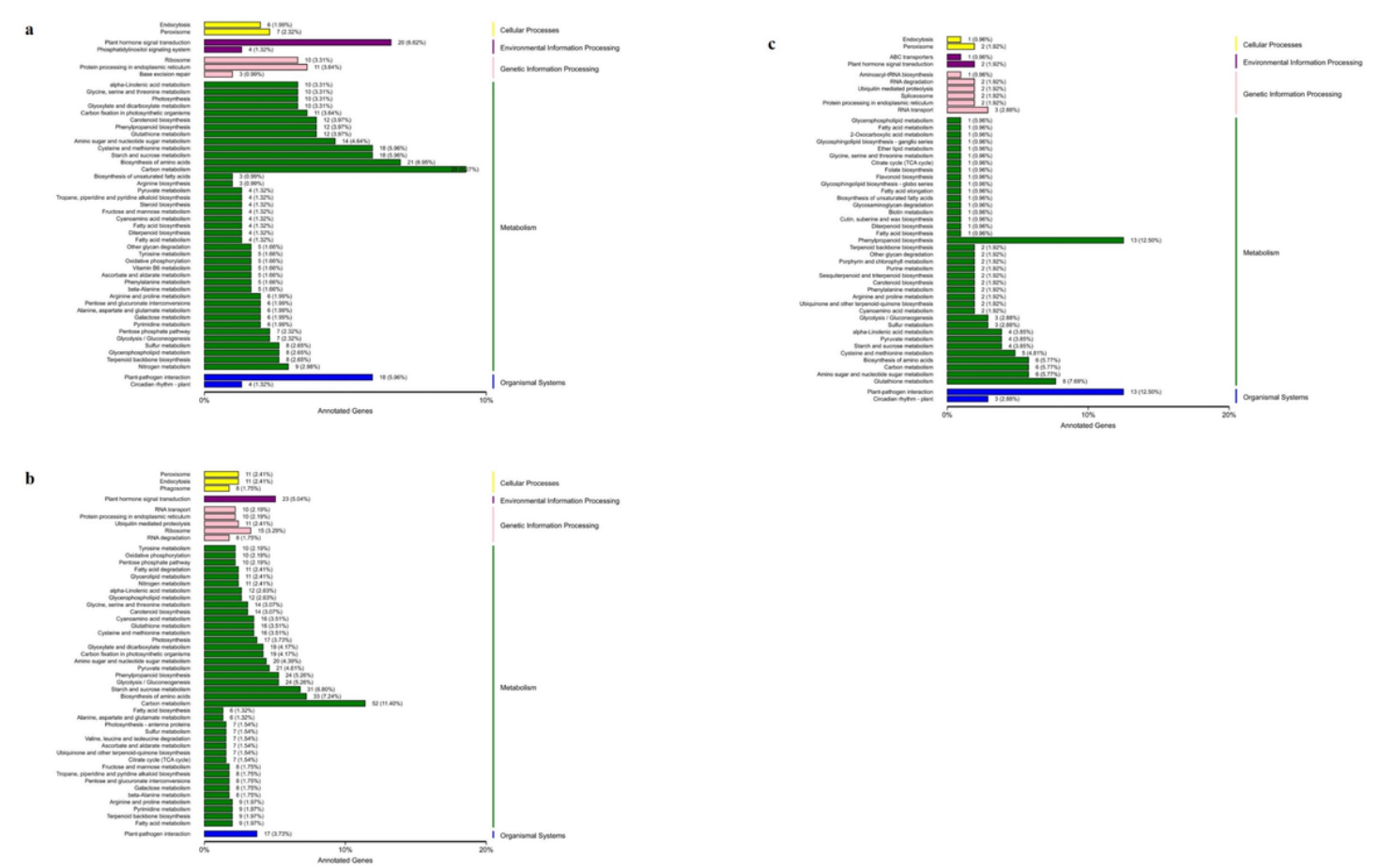


Figure 1

KEGG classification of differentially expressed genes in the different test materials. (a) KEGG classification of differentially expressed genes between BQ and DK. (b) KEGG classification of differentially expressed genes between BQ and ZJ. (c) KEGG classification of differentially expressed genes between DK and ZJ.

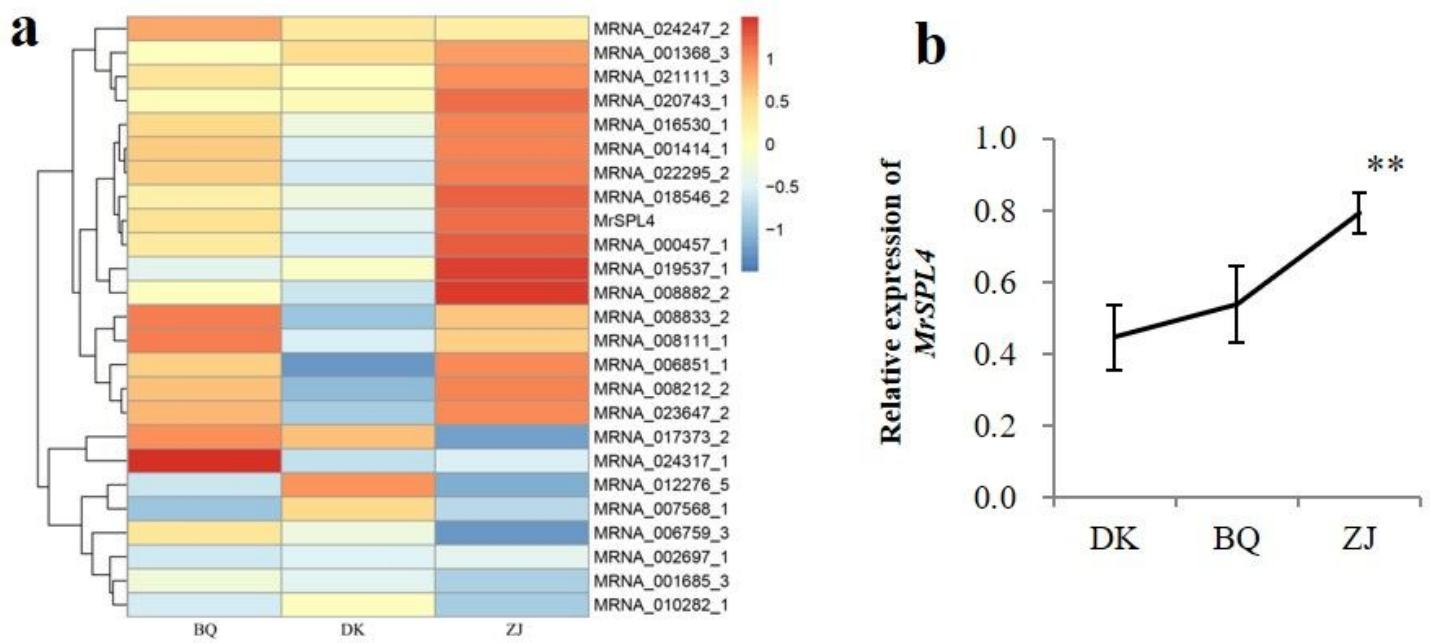


Figure 2

Expression of flowering-related differential genes. (a) The expression heat map of flowering related genes. (b) The relative expression of *MrSPL4* in different samples. Asterisks indicate values that are significantly different from those of the wild-type plants (** $0.001 < p < 0.05$).

Tree scale: 0.1

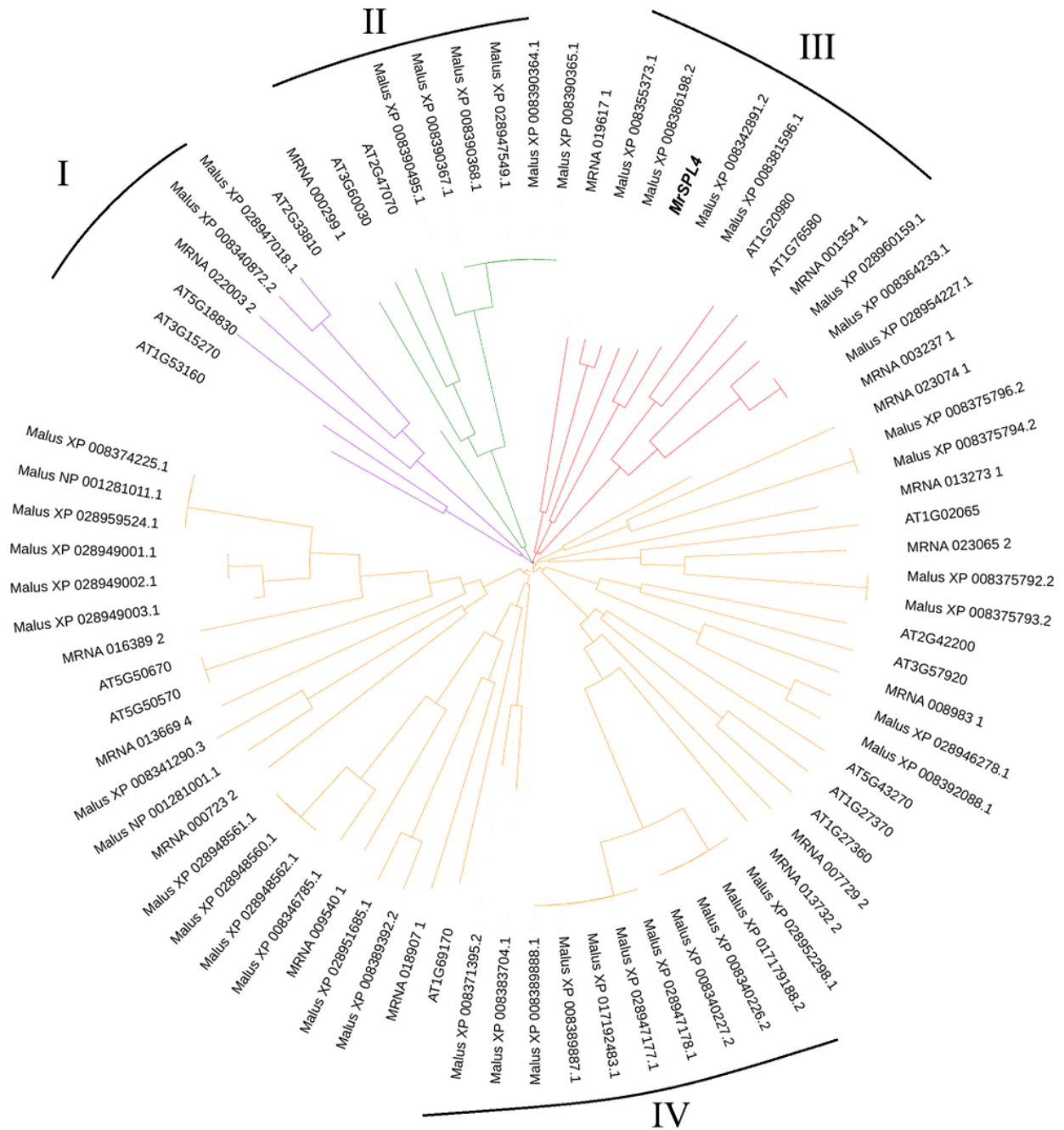


Figure 3

Evolutionary analysis of the SPL family gene in *Myrica rubra* and other species.

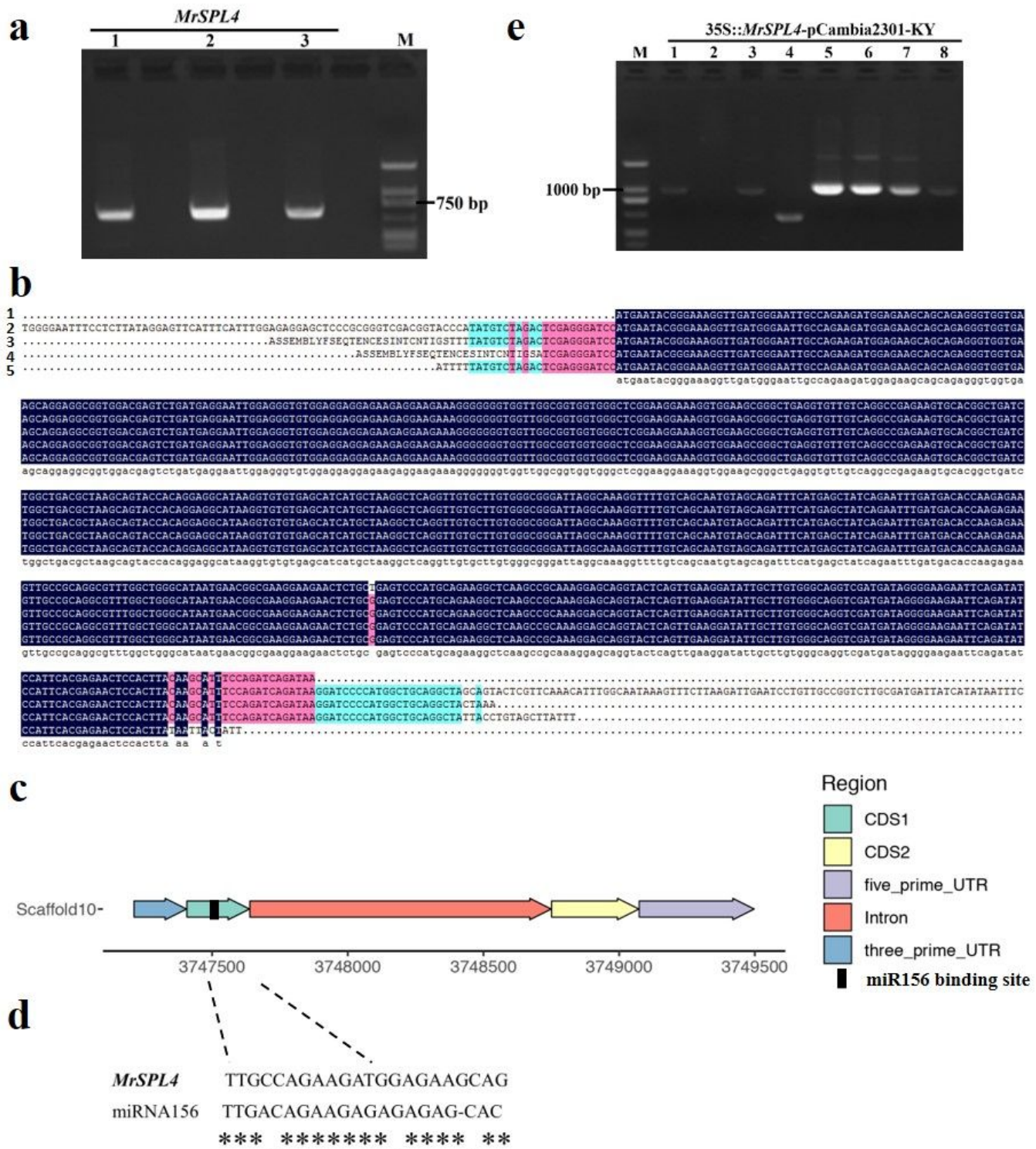


Figure 4

Cloning, sequence alignment, structure, mirna156 target and construction of overexpression vector of *MrSPL4* gene in different test materials. (a) The result of the coding region of *MrSPL4* electrophoresis, M: DL2000 DNA marker, 1-3: *MrSPL4* gene cloning in different samples. (b) The CDS sequence of *MrSPL4*, 1 refers to genome sequence, and 2, 3, 4 and 5 refer to gene sequence in ZJ, ZJ, BQ and DK, respectively. (c) The gene structure of *MrSPL4*. (d) The binding site of *miRNA156*. (e) The PCR identifying result of the

transformed bacterial solution, M: DL2000 DNA marker, 1-8: 35S::*MrSPL4*-pCambia2301-KY. Original blots/gels are presented in Supplementary Figure S1 2.

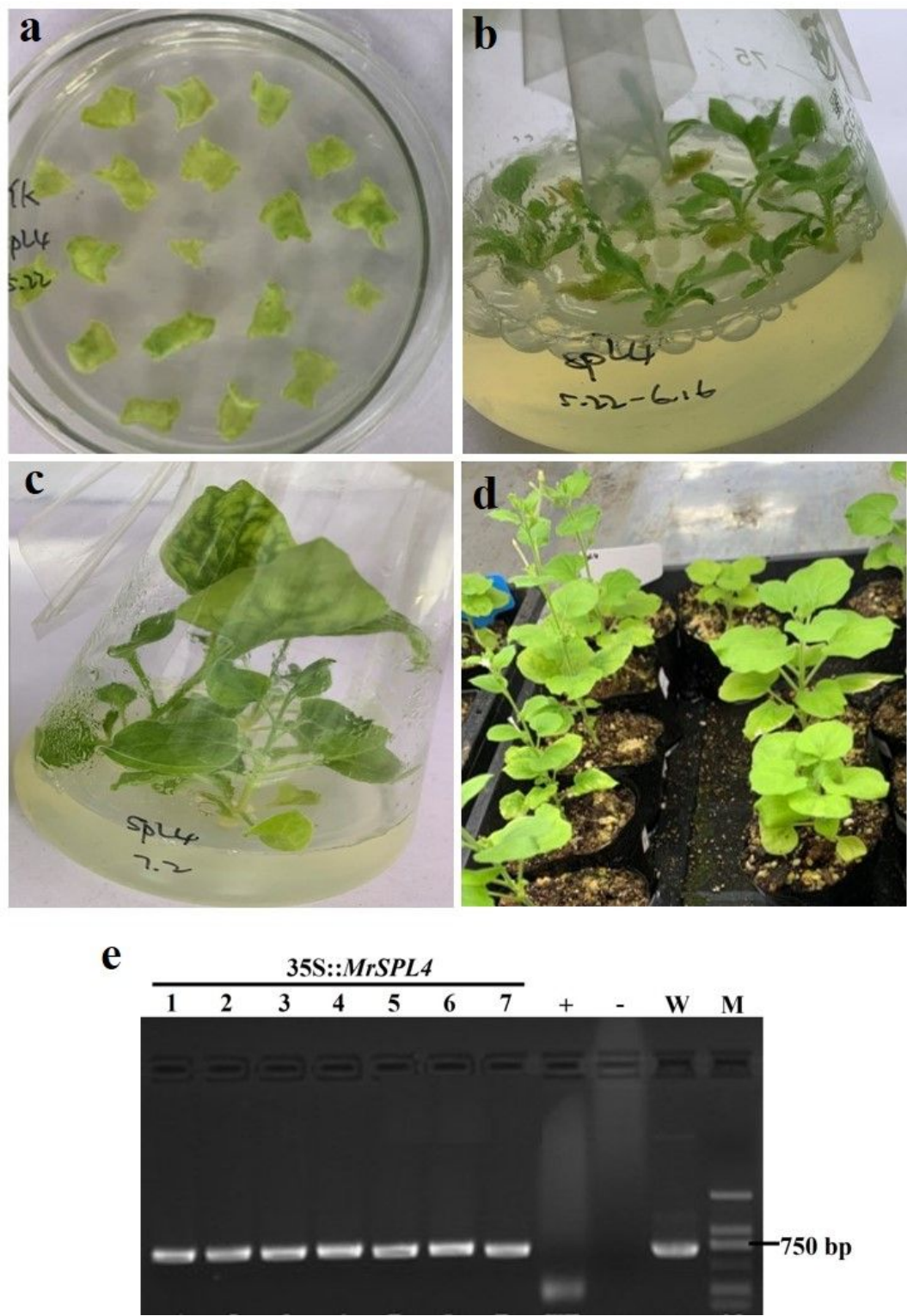


Figure 5

Regeneration of positive tobacco plants and PCR identification of Kan resistance. (a) Kanamycin medium for bud differentiation. (b) Resistant plant culture. (c) Rooting of resistant plants. (d) Transplanted

seedling culture. (e) Transgenic tobacco identification using genomic PCR, 1~7: resistant regenerated tobacco, W: negative control, -: blank control, +: positive control, M: DL2000 DNA marker. Original blots/gels are presented in Supplementary Figure S3.

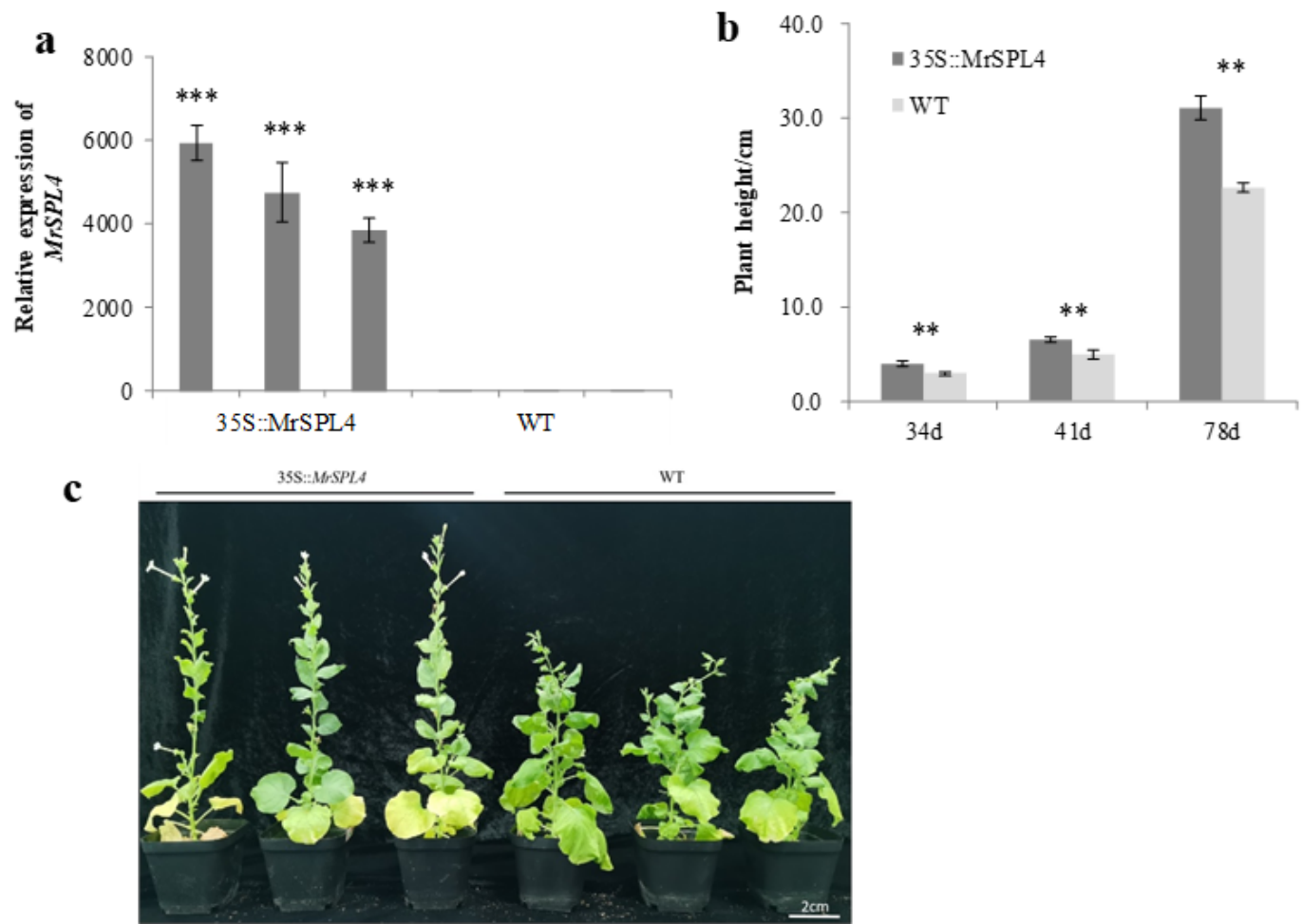


Figure 6

Analysis in the T₁ generation of positive tobacco lines. (a) *MrSPL4* expression in positive and WT tobacco. (b) The plant height of positive and WT tobacco. 34d, 41d and 78d indicate the number of days after transplanting. (c) The phenotype of positive and WT tobacco. Asterisks indicate values that are significantly different from those of the wild-type plants (** 0.001 < *p* < 0.05, *** *p* < 0.001).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformationfile.pdf](#)